

**INTERVAL SELISIH RATA-RATA DENGAN
METODE BOOTSTRAP PERSENTIL**

Akhmad Fauzy

Statistika FMIPA UII Yogyakarta & siswa Ph.D Statistika UPM Malaysia

N. A. Ibrahim, Isa Daud dan Mohd. Rizam Abubakar

Institut Penyelidikan Matematik (INSPEM) dan Jabatan Matematik
Universiti Putra Malaysia

Abstrak

We usually use traditional method to construct interval for the difference between two population means. This interval needs an assumption that sample is random and has normal distribution (symmetric). This article will discuss another method, namely bootstrap percentile. Bootstrap percentile method is more potential in constructing this interval and give shorter interval than the traditional method. This method does not need an assumption that the sample has to have a normal distribution.

Kata Kunci : difference means, interval, traditional method, bootstrap percentile

1. PENDAHULUAN

Suatu industri biasanya mempunyai lebih dari satu mesin produksi yang sama. Walaupun mesin-mesin produksi tersebut mempunyai spesifikasi yang sama, akan tetapi jumlah produksi yang dihasilkan dapat berbeda. Biasanya manajemen industri ingin melakukan perhitungan tentang selisih rata-rata produksi untuk masing-masing mesin.

Dalam menentukan nilai interval untuk selisih rata-rata biasanya digunakan metode tradisional. Rumus yang digunakan dalam metode tersebut memerlukan asumsi bahwa kedua sampel adalah random dan berdistribusi normal. Apabila dalam pengambilan kedua sampel tersebut random dan berdistribusi normal, maka sebenarnya kita dapat dengan mudah membuat interval selisih rata-ratanya (Walpole and Myers, 1978).

Telah diketahui bahwa metode bootstrap persentil adalah suatu metode berbasis komputer yang sangat potensial untuk dipergunakan pada masalah-

masalah ketidakbiasan dan keakurasian, khususnya dalam menentukan interval. Pada kondisi sampel random berdistribusi normal, perlu dicoba apakah metode bootstrap persentil masih yang terbaik dalam menduga interval selisih rata-rata (Efron and Tibshirani, 1993).

Tulisan ini bertujuan untuk menunjukkan bahwa interval selisih rata-rata pada sampel berdistribusi normal yang dihasilkan oleh metode bootstrap persentil akan lebih baik apabila dibandingkan dengan menggunakan metode tradisional seperti yang selama ini kita gunakan.

2. METODOLOGI PENELITIAN

Data yang digunakan pada penelitian ini adalah data buatan (*artificial*). Data tersebut menunjukkan jumlah produksi (dalam ton) yang dihasilkan oleh 2 mesin dengan spesifikasi yang sama selama 1 tahun. Langkah pertama adalah membuat pendugaan interval selisih rata-rata dengan metode tradisional. Langkah selanjutnya adalah mencari nilai ulangan bootstrap yang menunjukkan kondisi konvergen. Setelah diketahui kondisi konvergennya, maka interval dengan metode bootstrap persentil dapat dicari. Kemudian hasil pendugaan interval dengan metode tradisional dan metode bootstrap persentil dibandingkan.

2.1 Metode Tradisional

Dalam menentukan nilai interval yang dicari dengan menggunakan metode tradisional untuk selisih rata-rata pada sampel berdistribusi simetris dan variansi populasi tidak diketahui, para pengguna statistik biasanya menggunakan rumus $(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2} \sqrt{s_1^2/(n_1) + (s_2^2)/(n_2)}$ untuk jumlah sampel besar dan $(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2} \sqrt{s_1^2/(n_1) + (s_2^2)/(n_2)}$ untuk jumlah sampel kecil. \bar{x}_1, \bar{x}_2 dan s_1, s_2 masing-masing menyatakan rata-rata sampel dan deviasi standar sampel. Angka 1 dan 2 masing-masing menyatakan sampel kelompok 1 dan 2. $z_{\alpha/2}$ dan $t_{\alpha/2}$ menyatakan nilai dari distribusi z dan t dengan derajat bebas n_1+n_2-2 , dimana n_1 adalah jumlah sampel 1 dan n_2 adalah jumlah sampel 2. Rumus tersebut

memerlukan asumsi bahwa kedua sampel adalah random dan berdistribusi normal (Steel and Torrie, 1980).

Apabila dalam pengambilan kedua sampel tersebut random dan berdistribusi normal, maka sebenarnya kita dapat dengan mudah membuat interval selisih rata-ratanya. Akan tetapi apabila kedua sampel tersebut belum diketahui distribusinya, maka perlu diuji apakah kedua sampel berdistribusi normal. Apabila ternyata hasilnya tidak berdistribusi normal, maka biasanya kita terpaksa memberikan asumsi bahwa kedua sampel tersebut berdistribusi normal (Koopmans, 1987).

2.2 Metode Bootstrap Persentil

Metode bootstrap adalah suatu metode berbasis komputer yang sangat potensial untuk dipergunakan pada masalah ketidakstabilan dan keakurasian, khususnya dalam menentukan interval. Istilah bootstrap berasal dari “*pull oneself up by one’s bootstrap*”, yang berarti berpijak di atas kaki sendiri, berusaha dengan sumber daya minimal. Dalam sudut pandang statistika, sumber daya yang minimal adalah data yang sedikit, data yang menyimpang dari asumsi tertentu, atau data yang tidak mempunyai asumsi apapun tentang distribusi populasinya (Fauzy, 2000).

Tujuan dari penggunaan metode bootstrap adalah untuk mendapatkan pendugaan yang sebaik-baiknya yang berasal dari data yang minimal. Dengan demikian penggunaan komputer dalam metode bootstrap mutlak diperlukan (Fauzy, 1998).

Efron telah menunjukkan bahwa bootstrap dapat mengatasi berbagai persoalan yang muncul dalam pendugaan. Salah satu kendala dalam penerapan bootstrap pada awal penemuannya adalah penggunaan komputasi yang besar. Akan tetapi pada saat sekarang sejalan dengan perkembangan perangkat keras dan perangkat lunak komputer yang semakin canggih, penelitian dalam bidang

statistika yang berbasis komputer seperti bootstrap dapat berkembang sangat cepat (Zulela, et. al, 1997).

Secara umum prosedur bootstrap persentil untuk menduga interval selisih rata-rata adalah sebagai berikut :

1. Berikan peluang yang sama $1/n_1$ dan $1/n_2$ pada setiap pengamatan sampel berukuran n_1 dan n_2 ,
2. Mengambil suatu sampel berulang secara random berukuran n_1 dan n_2 dengan pengembalian, kemudian hitung selisih nilai rata-rata dari sampel 1 dengan sampel 2,
3. Ulangi langkah 2 sebanyak B kali untuk mendapatkan “*independent bootstrap replications*” $\hat{\beta}_n^{*1}, \hat{\beta}_n^{*2}, \dots, \hat{\beta}_n^{*B}$ dan mencari pada ulangan seberapa tercapai kondisi konvergen. Hitung selisih rata-ratanya, yaitu:

$$\hat{\beta}_n^* = \frac{1}{B} \sum_{i=1}^B (\hat{\beta}_{n_1}^{*i} - \hat{\beta}_{n_2}^{*i})$$

4. Interval bootstrap persentil pada tingkat $(1-\alpha)$ didefinisikan dengan persentil ke- $100(\alpha/2)$ dan ke- $100(1-\alpha/2)$ pada $\hat{\beta}_n^*$, sehingga interval persentil dapat dinyatakan dengan:

$$[\hat{\beta}_n^{*(\alpha/2)}, \hat{\beta}_n^{*(1-\alpha/2)}].$$

3. HASIL DAN PEMBAHASAN

Data yang digunakan pada penelitian ini adalah data buatan (*artificial*). Misalkan sebuah industri menghasilkan produk (dalam ton) yang dihasilkan oleh 2 mesin (X dan Y) dengan spesifikasi yang sama selama 1 tahun.

Tabel 1. Jumlah produksi (dalam ton) mesin X dan Y selama tahun 2002

bulan	mesin X	mesin Y	bulan	mesin X	mesin Y
Januari	563.8	379.0	Juli	534.4	477.8
Pebruari	494.7	427.6	Agustus	458.0	584.7
Maret	582.7	538.8	September	521.2	470.4
April	521.1	420.6	Oktober	468.1	532.3
Mei	578.4	433.9	Nopember	586.8	466.3
Juni	533.9	504.4	Desember	482.7	421.2

(data *artificial*)

Industri tersebut berkeinginan untuk mencari interval selisih rata-rata produksi antara mesin X dan Y, sehingga nantinya diperoleh gambaran tentang kapasitas dari mesin tersebut.

3.1 Metode Tradisional

Karena jumlah datanya kurang dari 30, maka untuk mencari interval selisih rata-ratanya digunakan rumus:

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2} \sqrt{s_1^2/(n_1) + (s_2^2)/(n_2)}$$

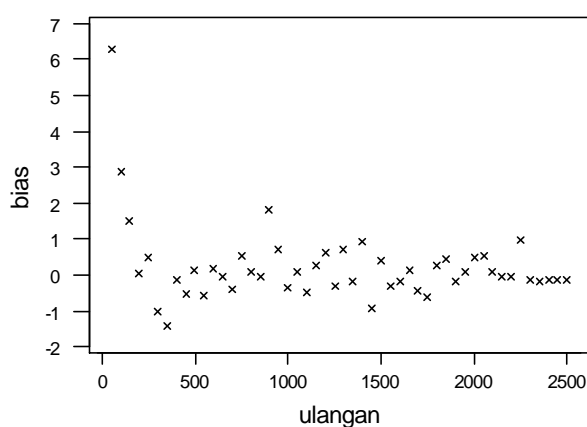
Dengan menggunakan rumus di atas maka batas bawah, batas atas dan lebar interval pada tingkat kepercayaan 99 % dan 95 % dapat diperoleh.

Tabel 2. Batas bawah (BB), batas atas (BA) dan lebar interval (LS) pada tingkat kepercayaan (TK) 99 % dan 95 % untuk selisih rata-rata produksi mesin X dan Y

TK	BB	BA	LS
99 %	(-14.1573)	108.9573	123.1146
95 %	2.1000	92.6902	90.5902

3.2 Metode Bootstrap Persentil

Untuk mencari pada ulangan seberapa akan tercapai kondisi konvergen, maka perlu dibuat plot antara bias dengan ulangan. Plot tersebut dapat dilihat pada grafik 1.



Grafik 1. Plot antara bias dengan ulangan pada selisih rata-rata

Dari grafik 1 terlihat bahwa kondisi konvergen tercapai pada ulangan ke-2300. Interval bootstrap persentil pada tingkat kepercayaan 99 % dan 95 % dapat dilihat pada tabel 3 di bawah ini.

Tabel 3. Batas bawah (BB), batas atas (BA) dan lebar interval (LS) pada tingkat kepercayaan (TK) 99 % dan 95 % untuk selisih rata-rata produksi mesin X dan Y

TK	BB	BA	LS
99 %	(-9.1667)	99.3917	108.5584
95 %	5.8583	88.3083	82.4500

3.3 Perbandingan Lebar Interval

Perbandingan lebar interval selisih rata-rata pada jumlah sampel kecil antara metode tradisional dengan metode bootstrap persentil dapat dilihat pada tabel 4.

Tabel 4. Lebar interval pada tingkat kepercayaan (TK) 99 % dan 95 % untuk selisih rata-rata produksi mesin X dan Y

Metode	TK	
	99 %	95 %
Tradisional	123.1156	90.5902
Bootstrap persentil	108.5584	82.4500
Selisih interval	14.5572	8.1402

Interval yang dihasilkan oleh metode bootstrap persentil lebih pendek dari pada yang dihasilkan oleh metode tradisional. Hal ini bisa dilihat pada tabel 4, dimana terjadi selisih lebar interval yang cukup besar antara kedua metode tersebut. Dengan demikian dapat disimpulkan bahwa metode bootstrap persentil jauh lebih baik dari pada metode tradisional, karena metode bootstrap persentil menghasilkan lebar interval yang lebih pendek.

4. KESIMPULAN

Metode yang lebih baik di dalam menduga interval selisih rata-rata pada sampel kecil adalah metode bootstrap persentil. Metode tersebut menghasilkan lebar interval yang lebih sempit apabila dibandingkan dengan metode tradisional.

DAFTAR PUSTAKA

1. Efron, B. and Tibshirani R.. *An Introduction to the Bootstrap*. New York: Chapman & Hall. 1993.
2. Fauzy, Akhmad.. Interval Konfidensi untuk Koefisien β_1 dari Garis Regresi apabila Ragam Galat Tidak Homogen dengan Metode OLS, WLS dan Bootstrap. Thesis. Bogor: IPB Bogor. Tidak dipublikasikan. 1998.
3. Fauzy, Akhmad. Estimasi Interval Konfidensi Nilai Rata-rata pada Sampel Berdistribusi t dengan Metode Bootstrap Persentil. Bandung: Jurnal MIHMI ITB, Volume 6 No. 5 ISSN 0854-1380 tahun 2000 Akreditasi DIKTI No. 2082/D/T/1996.
4. Koopmans, L. H. *Introduction to Contemporary Statistical Methods*. Second Edition. Boston: PWS Publisher 1987.
5. Steel, Robert G. D. and James H. Torrie.. *Principles and Procedures of Statistics*. New York: McGraw-Hill. 1980
6. Walpole, Ronald E. and Raymond H. Myers. *Probability and Statistics for Engineers and Scientists*. Second Edition. New York: Macmillan Publishing Co. 1978.
7. Zulaela, et. al. Simulasi Bootstrap untuk Estimasi Interval. Laporan Penelitian bidang Statistika. Lembaga Penelitian, Yogyakarta: Universitas Gadjah Mada. Tidak dipublikasikan. 1997.